|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tree Leaner Hold Out Validation  (Sem arrival) | Quality Measure | Pruning Method | Min number records per node | Number records to store for view | Score | Cohen’s Kappa(k) |
|  | Gain ratio | No pruning | 3 | 10000 | 94.999% | 0.898% |
|  | Gini index | No pruning | 3 | 10000 | 94.776% | 0.894% |
|  | Gain ratio | MDL | 3 | 10000 | 96.008% | 0.918% |
|  | Gini index | MDL | 3 | 10000 | 95.846% | 0.916% |
|  | Gain ratio | MDL | 5 | 10000 | 96.108% | 0.92% |
|  | Gain ratio | MDL | 10 | 10000 | 96.093% | 0.92% |
|  | Gain ratio | MDL | 7 | 10000 | 96.158% | 0.921% |
|  | Gain ratio | MDL | 8 | 10000 | 96.119% | 0.921% |
|  | Gain ratio | MDL | 6 | 10000 | 96.166% | 0.922% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tree Leaner Hold Out Validation  (Com arrival) | Quality Measure | Pruning Method | Min number records per node | Number records to store for view | Score | Cohen’s Kappa(k) |
|  | Gain ratio | No pruning | 3 | 10000 | 94.937% | 0.897% |
|  | Gini index | No pruning | 3 | 10000 | 94.876% | 0.896% |
|  | Gain ratio | MDL | 3 | 10000 | 95.95% | 0.917% |
|  | Gini index | MDL | 3 | 10000 | 95.838% | 0.915% |
|  | Gain ratio | MDL | 6 | 10000 | 96.077% | 0.92% |
|  | Gain ratio | MDL | 7 | 10000 | 96.046% | 0.919% |
|  | Gain ratio | MDL | 5 | 10000 | 96.023% | 0.919% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tree Leaner Cross Validation  (Sem arrival) | Quality Measure | Pruning Method | Min number records per node | Number records to store for view | Score | Cohen’s Kappa(k) |
|  | Gain ratio | No pruning | 3 | 10000 | 94.904% | 0.896% |
|  | Gini index | No pruning | 3 | 10000 | 94.828% | 0.895% |
|  | Gain ratio | MDL | 3 | 10000 | 95.91% | 0.917% |
|  | Gini index | MDL | 3 | 10000 | 95.908% | 0.916% |
|  | Gain ratio | MDL | 5 | 10000 | 95.986% | 0.918% |
|  | Gain ratio | MDL | 6 | 10000 | 96.008% | 0.918% |
|  | Gain ratio | MDL | 7 | 10000 | 96.006% | 0.918% |
|  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Naïve Bayes Hold Out Validation | Default Probability | Minimum Standard Deviation | Threshold standard deviation | Maximum number of unique nominal values per attribute | Score | Cohen’s Kappa(k) |
|  | 0.0001 | 0.0001 | 0.001 | 10 | 85.483% | 0.704% |
|  | 0.0003 | 0.0001 | 0.001 | 10 | 85.356% | 0.702% |
|  | 0.0001 | 0.0003 | 0.001 | 10 | 85.483% | 0.704% |
|  | 0.0001 | 0.0001 | 0.002 | 10 | 85.483% | 0.704% |
|  | 0.0001 | 0.0001 | 0.001 | 20 | 85.483% | 0.704% |
|  | 0.0003 | 0.0003 | 0.003 | 10 | 85.356% | 0.702% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Naïve Bayes Cross Validation | Default Probability | Minimum Standard Deviation | Threshold standard deviation | Maximum number of unique nominal values per attribute | Score | Cohen’s Kappa(k) |
|  | 0.0001 | 0.0001 | 0.00 | 20 | 85.326% | 0.701% |
|  | 0.0002 | 0.0001 | 0.00 | 20 | 85.333% | 0.701% |
|  | 0.0003 | 0.0001 | 0.0 | 20 | 85.205% | 0.699% |
|  | 0.0001 | 0.0002 | 0.0 | 20 | 85.326% | 0.701% |
|  | 0.0001 | 0.0001 | 0.0001 | 20 | 85.326% | 0.701% |

|  |  |  |  |
| --- | --- | --- | --- |
| Random Forest Hold Out Validation | Split Criterion | Score | Cohen’s Kappa(k) |
|  | Information Gain Ration | 96.451% | 0.927% |
|  | Information Gain | 96.377% | 0.926% |
|  | Gini Index | 96.316% | 0.925% |

|  |  |  |  |
| --- | --- | --- | --- |
| Random Forest Cross Validation | Split Criterion | Score | Cohen’s Kappa(k) |
|  | Information Gain Ration | 96.457% | 0.928% |
|  | Information Gain | 96.389% | 0.926% |
|  | Gini Index | 96.398% | 0.926% |

|  |  |  |  |
| --- | --- | --- | --- |
| Tree Ensemble Hold Out Validation | Split Criterion | Score | Cohen’s Kappa(k) |
|  | Information Gain Ration | 96.451% | 0.927% |
|  | Information Gain | 96.377% | 0.926% |
|  | Gini Index | 96.254% | 0.924% |

|  |  |  |  |
| --- | --- | --- | --- |
| Tree Ensemble Cross Validation | Split Criterion | Score | Cohen’s Kappa(k) |
|  | Information Gain Ration | 96.223% | 0.923% |
|  | Information Gain | 96.405% | 0.927% |
|  | Gini Index | 96.256% | 0.923% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Gradient Boosted Tree Hold Out Validation | Number of models | Learning rate | Score | Cohen’s Kappa(k) |
|  | 160 | 0.1 | 94.618% | 0.891% |
|  | 275 | 0.1 | 94.622% | 0.891% |
|  | 275 | 0.2 | 94.681% | 0.892% |
|  | 160 | 0.2 | 94.668% | 0.892% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Gradient Boosted Tree Cross Validation | Number of models | Learning rate | Score | Cohen’s Kappa(k) |
|  | 160 | 0.05 | 94.158% | 0.882% |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |